TECHNICAL REPORT 1995
December 2010

The Effects of Acute Stress on Cognitive Performance
A Pilot Study

Steven A. Murray
Matthew Yanagi
Wayne Ensign
SSC Pacific

Camellia Clark
Burcu Darst
University of California, San Diego

SSC Pacific

SSC PACIFIC
San Diego, California 92152-5001

J. J. Beel, CAPT, USN
Commanding Officer

C. A. Keeney
Technical Director

ADMINISTRATIVE INFORMATION

This report was prepared for prepared for the Office of Naval Research, Arlington, VA, by the Applied Research Branch, Space and Naval Warfare Systems Center Pacific, San Diego, CA.

Released by
G. W. Anderson, Head
Applied Research Branch

Under authority of
M. J. Machniak, Head
Advanced Systems and Applied
Sciences Division

LH

TECHNICAL REPORT 1995
December 2010

The Effects of Acute Stress on Cognitive Performance
A Pilot Study

Steven A. Murray
Matthew Yanagi
Wayne Ensign
SSC Pacific

Camellia Clark
Burcu Darst
University of California, San Diego

SSC Pacific
San Diego, CA 92152-5001

EXECUTIVE SUMMARY

A laboratory study was conducted to establish a research framework for investigating the effects of stress and fatigue on cognitive performance. The initial objectives were to (a) confirm the effectiveness of candidate stressor tasks, (b) to evaluate alternative stress response measures, and (c) to benchmark a series of cognitive performance tests. Both stressor tasks proved effective in eliciting stress under laboratory conditions, as indicated by multiple stress measures. All measures, however, were not effective, and no cognitive performance effects were found. Results are explained in terms of experiment design factors (i.e., the between-subjects approach used for the study) and the intensity and duration of stress levels achievable under laboratory conditions. Methodological revisions and an interim experiment are discussed in the context of larger research objectives that address both stress and fatigue together.

CONTENTS

Figures

Tables

INTRODUCTION

The effects of stress and fatigue on military performance have been well documented in both research and operational reports. Decrements to memory and logical reasoning (Leach, 2004; Lieberman, et al., 2005), decision making (Orosanu and Backer, 1996; Combat Stress Control, 1998), situation awareness (e.g., Sterling and Perala, 2007; Gorman, Cooke, and Winner, 2006), and learning (Joels, et al., 2006) are all observed in the context of operational stress. Similarly, the fatigue that accompanies the high operating tempos can negatively impact many components of cognitive performance, whether or not such operations would otherwise be considered “stressful.” Van Dongen, et al. (2003), for example, reported dose-response relationships between sleep restriction and performance decrements on structured vigilance and math tasks, while Habeck, et al. (2004) found a direct correlation between sleep deprivation, performance on a perceptual-memory task, and activation levels of related brain regions. Cognitive performance decrements in more complex decision making tasks have also been observed following varying levels of sleep deprivation (e.g., Harrison and Horne, 2000).

While stress and fatigue can each impact cognitive performance, no studies have been identified that address whether their contributions are independent or, if related, the nature of their interactions. In fact, because stress and fatigue typically occur together in both military and work settings the two factors are often discussed as a single construct (e.g., Davidson and Cooper, 1981; Raggatt and Morrissey, 1997; Friedl, et al., 2004).

An empirical distinction between stress and fatigue effects could enhance the power of current cognitive performance models and better focus both prevention and remediation methods for military and civilian work forces. Such a research program would require a structured investigation using a common set of response metrics and a common set of benchmark performance tasks. Additionally, because the consequences of acute versus chronic stress (e.g., McGonagle and Kessler, 1990) and acute versus chronic fatigue (e.g., Poteliakhoff, 1981) are likely different, a series of studies would likely be necessary to a complete understanding of these factors. The research reported here concerns the results of an initial effort to characterize and validate the cognitive effects of acute stress, as a preliminary step toward a full factorial investigation into both stress and fatigue.

ACUTE STRESS IN THE LABORATORY

A primary (but not the only) definition of stress is the mismatch between the demands of a task and the individual’s ability to cope with them (Lazarus and Folkman, 1984)—the greater the mismatch, the greater the stress. This definition implies that stress response might be manipulated to the degree that a mismatch can be established and controlled. Methods used for generating stress in research settings, include public speaking (e.g., the Trier Social Stress Test; Kirschbaum, Pirke, and Hellhammer, 1995), unpleasant environmental conditions (e.g., Brisswalter, Collardeau, and Rene, 2002), and challenging computer tasks (e.g., the Montreal Imaging Stress Task (MIST); Dedovic, et al., 2005). The Trier Social Stress Test requires a person to give a speech, before a group of judges to generate evaluative stress. Environmental conditions involve task performance in hot, cold, or noisy settings, or following physical exertion. The MIST presents a series of difficult math problems to be solved under time pressure. Stressful computer tasks are particularly attractive because they are inexpensive and portable, can be delivered with consistent control, and can be administered by a single person. Therefore, computer-based stress tasks are the focus of the study reported here.

MEASURING THE STRESS RESPONSE

Stress response signals can be detected with either psychological or physiological methods. An individual's psychological state can be determined by asking them to respond to either a direct query (e.g., interview) or to designate a description that most closely represents their personal state perception. Many self-report questionnaires, for example, require an individual to rate their current state (e.g., attitudes or mood, etc.) via a standardized checklist or set of short response items. The requirement for measuring an acute stress response, of course, is to select an instrument that responds to current (e.g., the Stanford Acute Stress Reaction Questionnaire (SASRQ); Cardena, et al., 2000), rather than more chronic (e.g., the Profile of Mood States (POMS); McNair, Lorr, and Droppelman, 1971) stress conditions.

Questionnaires are inexpensive and can be completed quickly—an important feature for gathering several measures in a short period of time—and most have good face and construct validity (i.e., appear to measure what they purport to measure). Data from questionnaires that are widely used (such as in research, clinical, or educational settings) can be accumulated over time to establish group and population norms, and can be benchmarked for validity and reliability. The State-Trait Anxiety Index (or STAI; Spielberger and Sydeman, 1994) is a popular self-report instrument of this type, and provides distinct scores for *state* anxiety (a property of the situation) and *trait* anxiety (a property of the individual) using a 4-point rating scale. The STAI requires a special scoring procedure before results can be interpreted, but has been effectively used for comparative studies of both anxiety and acute stress (e.g., Noto, et al., 2005; Chiffer McKay, et al., 2010). Other stress measurement questionnaires use a 5- or 10-point rating scale and can be evaluated without formal scoring (e.g., Kirschbaum, et al., 1995; Van Dongen, et al., 2004). The questionnaire described in Wang, et al. (2005), for example, contains 10-point scales for each of several stress-related dimensions—Stress, Anxiety, Effort, Frustration, and Difficulty—all on a single page.

While self-report (questionnaire) instruments are direct and useful, misunderstanding or misinterpretation of questionnaire items is always possible, and controls must be included to avoid intentional deception.

Psychological instruments can be <u>sufficient</u> to determine state (e.g., if the individual states that they're stressed, then they are) and are even appropriate to measure physiological conditions, a technique known as cross-modal matching, which queries the individual to report a physical state such as pain (e.g., Huskisson, 1982) or physical exertion (e.g., Borg, 1982) on a rating scale. Convergence between psychological and physiological measures represents a stronger inference basis for research, however, so a variety of measures is typically employed so that sensitivity and underlying generative mechanisms reflected by these approaches can be compared.

Physiological changes reflect body reactions to psychological or physical stimulation and, because physiological processes are generated internally (i.e., through neural or biochemical mechanisms), they are often interpreted as relatively independent of the consciously-mediated responses required by self-report questionnaires. Physiological methods are therefore used in human research as a substitute for, or complement to, psychological approaches. Common physiological performance metrics include cardiac function (such as heart rate and blood pressure; e.g., Vrijkotte, van Doornen, and de Geus; 2000), respiration (e.g., Grossman, 1983), electrical skin conductance (e.g., Horvath, 1978) and analysis of stress hormones in the blood, saliva, or urine (such as cortisol or human nerve growth factor; e.g., Kirschbaum and Hellhammer, 1999; Steptoe, Hamer, and Chida, 2007; Aloe, Alleva, and Fiore, 2002). Of these, cardiac function offers considerable research precedent for evaluating a range of stress values. Furthermore, salivary hormone sampling is relatively quick and easy to perform, and requires almost no equipment. Together, cardiac and hormone methods capture both biochemical and neurological phenomena.

The primary disadvantage of physiological measures is that the body processes upon which they are based are influenced by many factors besides the stimulus of interest. Dietary and drug habits, physical activity (even talking), state of health, time of day, etc. can dramatically alter physiological indices. Interpretation of physiological measures is also complex, as different mechanisms control different processes. Cortisol levels, for example, are controlled primarily by the hypothalamic-pituitary-adrenal axis (HPA) while nerve growth factor (NGF) levels are controlled primarily by the amygdala-medullary axis that, in turn, modulates the HPA (e.g., Aloe, et al., 1986).

The exploratory purpose of the study reported here dictates a multivariate approach to measuring the stress response, to ensure that the reactions generated by the stressor tasks are fully characterized. Both psychological (self-report) and physiological measures will therefore be employed.

MEASURING COGNITIVE PERFORMANCE

A wide variety of cognitive performance measures exists, derived from research, clinical, industrial and educational sources. Tests are available for both general applications (e.g., the Psychomotor Vigilance Test (PVT), which measures sustained response time; Thorne, et al., 2005) and specialized requirements (e.g., dementia screening; Mioshi, et al., 2006). While some cognitive tests can be administered with paper-and-pencil, the need to evaluate the speed of psychomotor processing typically requires computer administration of timed stimuli. Selection of measurement tools for the current study is driven primarily by the expected impact of stress and fatigue on specific cognitive characteristics. That is, while the current phase of the study is focused on acute stress, the same measurement tools must also be relevant to later phases that will include fatigue conditions. Typical cognitive performance decrements associated with these factors, extracted from the research literature, may be summarized as follows:

<u>Acute Stress</u>

- Memory (e.g., Kirschbaum, et al., 1996; Vedhara, et al., 2000)
- Logical reasoning (e.g., Leach, 2004; Lieberman, et al., 2005)
- Decision making (e.g., Keinan, 1987; Orosanu and Backer, 1996)
- Learning (e.g., Yehuda, et al., 1995; Joels, et al., 2006)

<u>Fatigue</u>

- Vigilance (e.g., Krueger, 1989; Van Dongen, et al., 2003)
- Math processing (e.g., Wang, et al., 2005; Gunzelmann, et al., 2007)
- Perceptual processes (e.g., Krueger, 1989; Habek, et al., 2004)
- Decision making (e.g., Rosekind, et al., 1994; Harrison and Horne, 2000)

In fact, stress and fatigue effects are not orthogonal and examples of each decrement category, above, can be found in both types of research literature. The range of effects is significant, however, which means that several approaches to cognitive performance testing are necessary if stress and fatigue effects are to be characterized and—more importantly for the current research—distinguished from one another. The Automated Neuropsychological Assessment Metrics (ANAM; Reeves and Winter, 1992) is somewhat unique among cognitive evaluation tools in providing just such a diverse set of component tests. Furthermore, sub-tests can be selected from the full battery to suit specific needs, which provides considerable flexibility for focused research applications.

RESEARCH OBJECTIVES

The current study represents a foundation step for the larger topic of investigating the independent and combined effects of acute stress and fatigue on cognitive performance, and is therefore, exploratory. The precursor objectives addressed in the work reported here are to:

1. Establish that a measurable stress response can be generated in the laboratory among members of a general population. Specifically, the goal is to use challenging computer tasks to induce feelings of performance failure or inadequacy (i.e., a mismatch between demand and coping ability).
2. Determine whether or not deception (e.g., additional interactions with subjects to artificially exacerbate their feelings of performance failure) is necessary to achieve a significant stress response.
3. Select a single computer task for the alternatives of this study, for future research phases.
4. Evaluate the impact of the stress response on multiple cognitive performance characteristics.
5. Compare the sensitivity and consistency of multiple stress response measures in the context of the laboratory stress setting, with previous research literature and with each other. Future phases of this work will utilize only the most diagnostic of these measures.

METHOD

The logic of the experiment was to gather cognitive performance data while the individual is presumably in an elevated stress state. The study protocol therefore involves a fixed temporal pattern—resting baseline, stress, performance testing and recovery—with stress measurements collected throughout. A between-subjects approach (i.e., comparing performance across groups) was employed to preclude any learning effects from multiple exposures to either the stressor task or to the cognitive performance tests.

STRESS TASKS

Two computer-based stress tasks were selected for use in the study: the <u>Montreal Imaging Stress Test (MIST)</u> and <u>Virtual Battlespace 2 (VBS2)</u>.

The MIST requires an individual to solve paced arithmetic problems without pencil, paper, or calculator and to designate answers with keyboard selections (see Dedovic, et al., 2005, for a complete description). MIST software measures the accuracy of answers in real time, and increases problem difficulty as necessary to maintain a desired level of task difficulty.

VBS2 is a combat simulation system that can be presented on a laptop computer (see Bohemia Interactive, 2010 for a complete description). The software is used by the U.S. Marine Corps and other services for operational training and has considerable capabilities beyond the applications of this research, including distributed exercises by multiple military units. The experiment task involved a "first person shooter" scenario with the user in the role of an infantry soldier navigating a street in an Iraqi city. The scenario required the user to defend against insurgent threats (such as snipers, suicide bombers, and improvised explosive devices) that were embedded within the local population and geography.

The MIST or the VBS2 task were administered to subjects based on random assignment, and response measures were compared to a control group that received no stress task at all. Both tasks were used in this initial experiment design for comparison purposes, i.e., (a) to ensure that a measurable stress response was obtained from both tasks and (b) to determine that the resulting stress responses were roughly equivalent. Presuming adequate performance, the VBS2 task would be retained for future research owing to its closer relevance to military operations.

Deception Condition

Each stress task group (i.e., MIST and VBS2) was further divided into a <u>stress-only</u> cohort that simply completed the task, and a <u>deception</u> cohort that received additional experimenter interaction during the task. Specifically, each member of the deception cohort was told that their performance was substandard between the first and second task sessions and, again, following the second session. The intent of deception was to amplify the stress response with additional social pressure, and to determine if such elevated response had an effect on the cognitive performance tests that immediately followed the stress manipulation. Presuming that adequate stress response was obtained without deception, then deception would not be included in future research phases.

STRESS MEASURES

The stress state of each subject was measured at pre-determined points during the test session, using a variety of methods, as follows:

Psychological (Self-Report) Measures

- The 20-item version of the State-Trait Anxiety Index (STAI), which requires a special scoring protocol to yield a state anxiety rating.
- A multi-factor Stress Scale (described in Wang, et al., 2005). This instrument was selected because it is short, involving 10-point Likert ratings for each of five stress dimensions—Stress, Anxiety, Effort, Frustration, and Task Difficulty—and direct (i.e., the questionnaire asks respondents how they feel for each of the stress dimensions, without the need for coded scoring procedures). The multi-dimensional nature of the Stress Scale may, furthermore, provide a more nuanced characterization of the stress response than a unitary score.

Physiological Measures

- Heart rate (HR), representing the five-minute average of beats per minute (BPM) prior to critical events in the experiment timeline. Increased stress is typically accompanied by an elevated heart rate (e.g., Vrijkotte, van Doornen, and de Geus, 2000).
- Heart rate variability (HRV), represented by the five-minute average of the ratio of low frequency to high frequency spectral power, taken prior to critical events in the experiment timeline. Because increased stress is associated with reduced parasympathetic (low frequency) activity and greater sympathetic (high frequency) activity, this ratio is expected to diminish (e.g., Filaire, et al., 2009).
- Salivary cortisol, a steroid hormone produced by the adrenal gland. Elevated cortisol levels have been found in human blood and saliva in response to stress, although time delays of several minutes post exposure have been observed (e.g., Wang, et al., 2005).
- Salivary Nerve Growth Factor (NGF), a protein molecule associated with growth of sympathetic and certain sensory nerves. Elevated NGF levels have been found in humans and animals in response to stress (e.g., Aloe, et al., 2002.

Hormone components were separately assayed from a common set of saliva samples, gathered with sublingual lozenges and salivettes. The collection procedure was standardized as recommended by Salimetrics, LLC (www.salimetrics.com). Samples were stored in a freezer immediately following collection and batch shipped to Salimetrics on dry ice for analysis. Cardiac measures were extracted from a continuous data record collected with an Aria Holter Digital Recorder (delmar Reynolds; see Medcompare, 2010, for current information), using a four-electrode configuration on the subject's chest, and worn throughout the experiment session.

COGNITIVE PERFORMANCE MEASURES

A test series selected from the Automated Neuropsychological Assessment Metrics (ANAM®, version 4.0) battery was administered immediately following the stress manipulation to identify decrements in the cognitive capabilities described earlier, i.e.:

- Memory
- Logical reasoning
- Decision making
- Learning
- Vigilance
- Math processing
- Perceptual processes
- Decision making

The selected ANAM battery consisted of the following tests (see C-SHOP, 2010 for complete task descriptions and graphics):

1. Modified Stanford Sleepiness Scale. Provides a self-assessment of sleep / fatigue state.
2. Mood Scale II – Revised. Provides a self-assessment of mood state in terms of seven categories: Vigor, Happiness, Depression, Anger, Fatigue, Anxiety, and Restlessness.
3. Code Substitution – Learning. Requires the subject to designate whether a symbol pair matches a preset list of pairings or not. *Tests visual search, sustained attention, and working memory*.
4. Match to Sample. Requires the subject to determine whether a spatial configuration matches another configuration, following a time delay. *Assesses spatial processing and visuo-spatial working memory*.
5. Logical Relations. Requires the subject to determine whether a verbal relation and a symbolic relation are matched or not. *Assesses abstract reasoning and verbal syntax ability*.
6. Mathematical Processing. Requires the subject to calculate and classify the result of a math problem. *Tests computational skills, concentration, and working memory*.
7. Visual Vigilance. Requires the subject to react quickly to an intermittent target, appearing at random intervals and positions on a blank display. *Tests sustained attention*.
8. Code Substitution – Delayed Memory. Requires the subject to respond to Code Substitution task without access to the standard pair display (i.e., the subject must rely on memory of pairings presented earlier to determine whether the current pairing matches the master list or not). *Assesses memory processes*.
9. Two-Choice Reaction Time. Requires the subject to react to only one of two randomly presented symbols. *Assesses choice reaction time*.
10. Four-Choice Reaction Time. Requires the subject to place a cursor over a spatially-randomized symbol as quickly as possible. *Assesses visuo-spatial processing*.
11. Simple Reaction Time. Requires the subject to respond to a target, appearing at random intervals, as quickly as possible (Similar to the Psychomotor Vigilance Test; PVT). *Assesses simple reaction time*.

The ANAM tests were always administered in the order above. With the exception of the Sleepiness and Mood scales, each test consisted of an instruction phase, a practice phase, and the actual test phase (i.e., where performance was measured). Each ANAM performance test provided a response time (RT) score, an accuracy (Percent Correct) score.

The transition between tests, and between each of the three test phases (above), was controlled by the subject via a keyboard entry, which allowed the entire ANAM battery to proceed with little or no experimenter intervention. The cognitive performance phase of the protocol was, however, interrupted following completion of the Mathematical Processing task to allow time for collection of stress measures, and then resumed to completion.

DESIGN

The factors of the experiment included:

- A two-level task factor, involving MIST and VBS2.
- A two-level stress factor involving participants who completed the stress task, and other participants who did not (i.e., a civilian Control group).

- A two-level deception factor for social stress (involving only those who completed a stress task).
- A control condition, involving all stress measurement collections and a final ANAM battery, but no stress task; subjects were free to read during the period normally allotted to the stress task.

The general design is given in Table 1.

Table 1. Experiment Design.

Stress				Control
MIST		VBS2		
No Deception	Deception	No Deception	Deception	

SUBJECTS

Subjects were recruited from open advertisements on local college campuses and in local news outlets. Requirements for participants were based primarily on the need to control for external influences on alertness and diurnal hormonal cycles, and included:

- Males
- Age 18–30
- Sufficient rest in the previous 24 hours
- No medications (evaluated on a case-by-case basis)
- Non-tobacco users
- No caffeine on the testing day

Test Environment

Each subject was tested individually. The experiment environment included a workstation with the stress task computer, instruction placards, salivette collection tubes, headphones, and an electronic timer. The experiment was administered by two researchers who administered instructions, delivered stimuli, and monitored progress from behind the subject, as shown in Figure 1.

Figure 1. Experiment Layout.

PROCEDURE

All individuals who responded to the public solicitation were first screened via telephone interview to ensure that basic participation requirements (above) were met. Those accepted into the study were provided with an appointment date and written instructions regarding pre-experiment sleep, exercise, and caffeine constraints. On the testing date, the following sequence of steps was administered to the subject:

1. Informed Consent procedures were completed and an orientation to the experiment was provided by one of the researchers. The experimenter explained the purpose of the study and the involvement of difficult or stressful task exposure.
2. The subject was asked about their current state of general health and recent sleep status.
3. The subject was asked to rinse their mouth by drinking water, to ensure better precision with saliva sampling.
4. A (*Baseline 1*) saliva sample was collected, using commercial (Salimetrics, LLC) salivette tubes and standard collection procedures. The collection task required the participant to soak a small synthetic fiber lozenge under their tongue for 90 seconds, and then to spit it into a test tube. The experimenter provided instructions prior to sample collection, and ensured the proper soaking period with an electronic timer. Immediately following collection, all salivettes were stored in a freezer.
5. An Aria heart rate monitor was placed on the subject, using a a 4-lead chest configuration. A marking button was pressed on the Aria to initiate cardiac and elapsed time recording.
6. A second (*Baseline 2*) saliva sample was collected to assess changes in stress hormones due to wearing the heart monitor.
7. The subject then completed a set of paper-and-pencil surveys, including"
    - A listing of food intake during the testing day (to assess caffeine and sugar intake)
    - The <u>Questionnaire of Competence and Control</u> (QCC; Krampen, 1991). A set of 32 statements requiring ratings of agreement on a six-point scale, saved for later analysis of possible personality factors in the data.
    - The <u>Rosenberg Self-Esteem Scale</u> (RSE; Rosenberg, 1979). A set of ten statements requiring ratings of agreement on a four-point scale, saved for later analysis of possible personality factors in the data.
8. The experimenter then provided instruction and experience with control procedures for the stress task by reading from a prepared script. One of three assignment conditions was possible for a subject, based on random selection—MIST, VBS2, or no stress task (i.e., Control). Stress task training was standardized by requiring the experimenter to read and demonstrate from a written script, and involved five minutes of hands-on execution. The experimenter commented on subject performance only for instructional purposes, and all questions were answered during and following training. Control subjects were given no instructions, but were allowed to read either their own materials or magazines provided by the experimenters. No reason was provided to control subjects for this, and the presumption was that they were to relax while the experiment was being prepared. *Subjects wore sound suppressing headphones during all task activities to reduce distractions*.
9. Following the five-minute training session, a third (*Training*) saliva sample was collected, for later comparison with the *Baseline* samples, to identify any elevation of hormones following the task training experience.
10. The subject was then asked to complete the State-Trait Anxiety Index and Stress Scale (*Training*)

11. The first of two stress tasks was administered, which involved a six-minute exposure to either the MIST or VBS2 tasks. The experimenter noted the event times on the Aria recorder by pressing the system button just before and, again, immediately after the stress task. Control group subjects were allowed to continue reading during this time, with no Aria mark or other task requirements.
12. At the conclusion of the first stress task administration, subjects in Deception groups were evaluated by the experimenter, who commented verbally that their performance was surprisingly poor and that they should try to do better during the next (second) task session. Subjects in the standard stress groups received no interaction.
13. Subjects then completed a second Stress Scale (*Stress 1*), to compare with the pre-stress baseline administration of step 10.
14. The stress task—MIST or VBS2—was administered for a second six-minute session. Control subjects continued to read or otherwise occupy themselves during this time. *Pre- and post-session event times were recorded using the Aria marker button*.
15. Following the stress task, subjects in the Deception group were again informed by the experimenter that their performance had been exceptionally poor. Subjects were told that they would have to complete the stress task again, at the end of the normal session, but that other tasks had to be completed next to stay within the required experiment timeline. No such interaction occurred for the standard stress subject groups or for the control group.
16. Another saliva sample, STAI, and Stress Scale (*Stress 2*) were collected at this time, i.e., when stress was presumably at its highest level for those who had completed a stress task.
17. All subjects were then instructed in the procedures required to complete the ANAM battery, and were permitted to begin when ready. The test sequenced was divided approximately into two equal parts (with a break following the Math task). Although the ANAM program provided for independent task completion and pacing by the subject, experimenters observed performance from the rear of the testing room and provided corrective feedback if the subject appeared to have misunderstood task requirements. *Event times were recorded on the Aria at the beginning and end of the ANAM set*.
18. At the break in the ANAM series, a fifth (*ANAM 1*) cortisol sample was collected.
19. The subject was then directed to complete the ANAM series. Beginning and concluding event times were recorded on the Aria.
20. When the ANAM battery was finished, another cortisol sample, STAI, and Stress Scale (*ANAM 2*) were collected.
21. At this point, the subject was debriefed by the experimenter using a standardized written script, including the intentional difficulty of the stress task, the purpose of experimenter interaction (Deception) and / or the need for control conditions, as appropriate.
22. Following a 15-minute delay period to allow the participant to return to a resting state, a final (*Recovery*) saliva collection was completed and the Aria heart monitor was removed.
23. After the final sample collection, the experimenters answered any remaining questions, provided payment, and released the subject.

A timeline of the experiment procedure, which lasted approximately 2.5 hours, is shown in Figure 2.

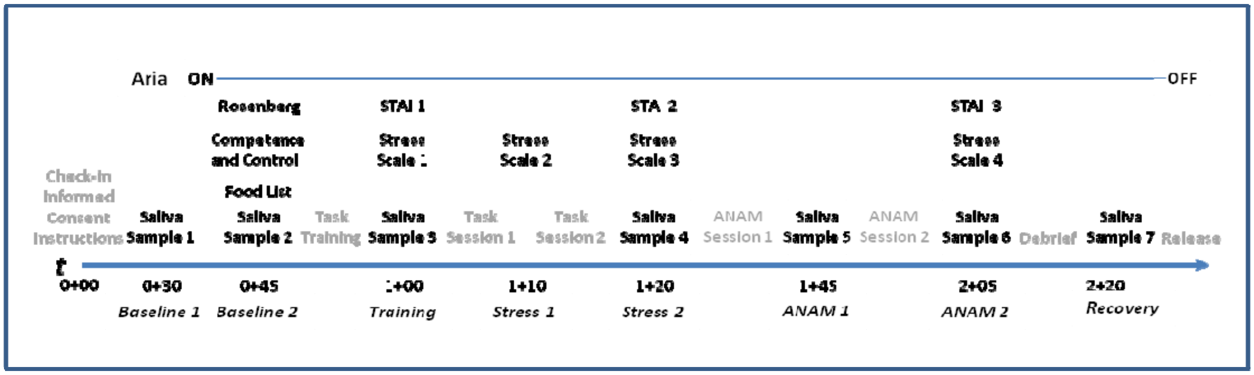

Figure 2. Experiment Procedure Timeline.

ANALYSIS

STAI and Stress Scale forms were scored by hand, using standardized rating procedures. Salimetrics, LLC assayed each saliva sample in duplicate for cortisol and in triplicate for NGF. Results were delivered to the experimenters as a spreadsheet data table. Cardiac data records were analyzed by the Laboratory for the Study of Emotion and Cognition Dr. Lilianne Mujica-Parodi, Director), Department of Biomedical Engineering, Stony Brook University, NY. Data series were analyzed in five-minute segments, shifted in one-minute increments, to generate time averages for heart rate (HR) and spectral power measures for HRV.

Final data were submitted to ANOVA processing, using STATISTICA 7 software (Statsoft, 2010), to determine significant differences between groups for both stress measures and cognitive performance measures, using a threshold level of $p \leq .05$ for statistical significance.

RESULTS

The study was designed to include 20 subjects in each of the two stress task conditions (i.e., 40 total subjects) and 20 subjects in the control condition. Half of each stress condition (10 subjects) was designated to receive the deception treatment involving additional experimenter interaction. All subject assignments were made with a blind rotation system, i.e., each condition received succeeding subjects in a fixed order, until all required subjects had been tested.

A total of 67 subjects were tested: 21 MIST subjects (including 11 in the stress-plus-deception condition), 22 VBS2 subjects (including 10 in the stress-plus-deception condition), and 21 control subjects. One subject withdrew and two subjects yielded unusable data.

SELECTION OF RESPONSE MEASURES

The set of response measures was first examined for inter-relationships, to select the smallest number of useful measures for analysis. A Spearman correlation matrix of Stress Scale component measures is presented in Table 2. These correlations were calculated for all participant groups and all data, including the stress *Recovery* data point, and significant ($p \leq .05$) correlations are rendered in bold print. Based on the many significant correlations among the sub-scales, only the STRESS component was used for subsequent performance measurement.

Table 2. Spearman Correlation Matrix – Stress Scale Components.

	STRESS	ANXIETY	EFFORT	FRUSTRATION	DIFFICULTY
STRESS		0.873	0.372	0.828	0.739
ANXIETY	0.873		0.347	0.82	0.670
EFFORT	0.372	0.347		0.420	0.441
FRUSTRATION	0.828	0.816	0.420		0.758
DIFFICULTY	0.739	0.670	0.441	0.758	

Table 3 shows another comparison of remaining stress response measures, with significant correlations ($p \leq .05$) rendered in bold print. With the exception of STAI and STRESS, none of these correlations was sufficiently large to exclude as a performance measure, owing to the possibility that each might be measuring a relatively unique aspect of stress. Despite the high correlation between STAI and STRESS, both were retained as the Stress Scale has not been validated in the research literature, and depending on a single psychological scale was deemed unwise.

Table 3. Spearman Correlation Matrix – Stress Response Measures.

	STAI	STRESS	Heart Rate	HRV	Cortisol	NGF
STAI		0.814	0.197	-0.055	0.034	0.114
STRESS	0.814		0.173	-0.042	0.053	0.191
Heart Rate	0.197	0.173		0.261	0.310	0.107
HRV	-0.055	-0.042	0.261		-0.052	-0.075
Cortisol	0.034	0.053	0.310	-0.052		0.129
NGF	0.114	0.191	0.107	-0.075	0.129	

OBJECTIVE A – CONFIRMING THE STRESS RESPONSE

The primary goal of the experiment was to determine whether a measurable stress response could be elicited by each stressor task. The time patterns of this response are shown in Figure 3 for each of the selected response measures. These results compare the (non-stress) Control group subjects with all Stress group subjects (i.e., both Stress and Stress-plus-deception).

The figure shows generally satisfactory results for STAI and STRESS (i.e., the two psychological instruments), but more complex patterns for the physiology measures. Elevated heart rate (HR), cortisol, and nerve growth factor (NGF), and depressed HRV, would be the expected responses to stress, based on previous literature. Although these patterns can be discerned for HRF and NGF they are, at best, difficult to detect for HR and cortisol.

The cortisol result is easier to identify in Figure 4, which depicts the natural logarithm transformation of these data, a process that serves to stabilize high variability (e.g., Box, Hunter and Hunter, 1978). A natural logarithm transformation for NGF data is also shown using this conversion, to make the raw result of Figure 3 easier to discern.

A gradually decreasing stress response can be observed for the Stress groups during ANAM administration, i.e., following cessation of the stress task manipulation. Conversely, some elevation in stress response can be observed in the Control group at the final ANAM test point, as the ANAM task—which all subjects completed—represented a relative increase in task stimulation level for this group.

The effectiveness of the computer-based tasks to elicit a stress response was evaluated statistically by comparing the Control subjects (who received no task) with all Stress subjects (i.e., including both tasks and the Deception condition). A summary of significant one-way ANOVA tests for this comparison is shown in Figure 5. Results for both psychological instruments and two physiological measures (HRV and NGF) were significant, and in the expected direction.

While the stress measurement tools were not consistently significant, those that were provide convergent evidence that the stress tasks were successful in generating stress with a laboratory task.

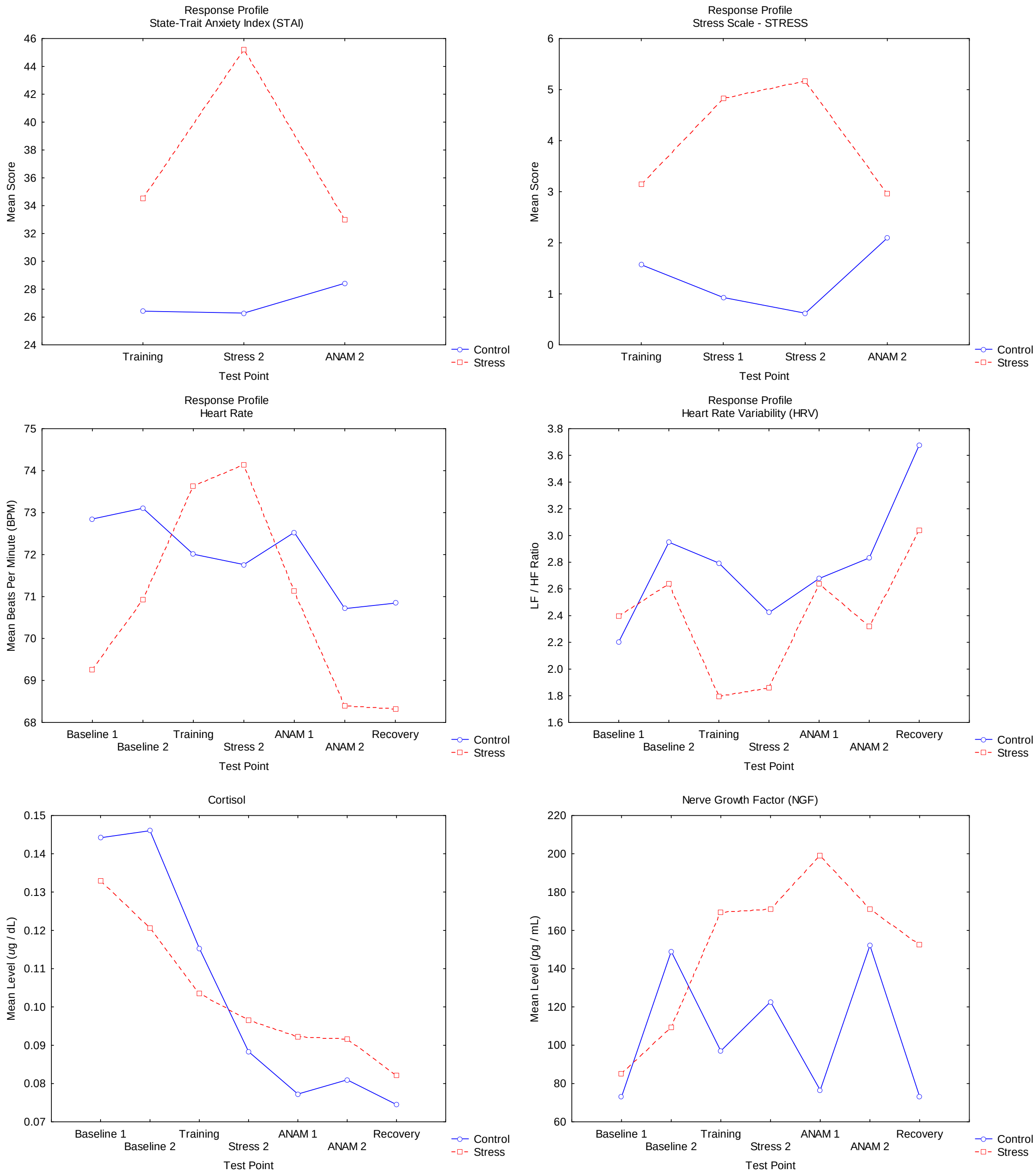

Figure 3. Time Profiles – Selected Stress Response Measures.

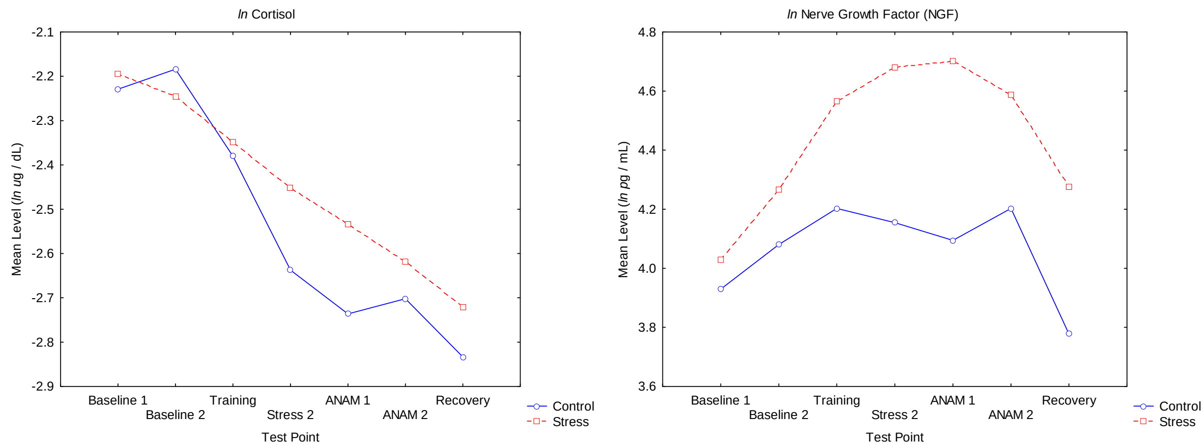

Figure 4. Time Profiles – Natural Logarithm Conversions.

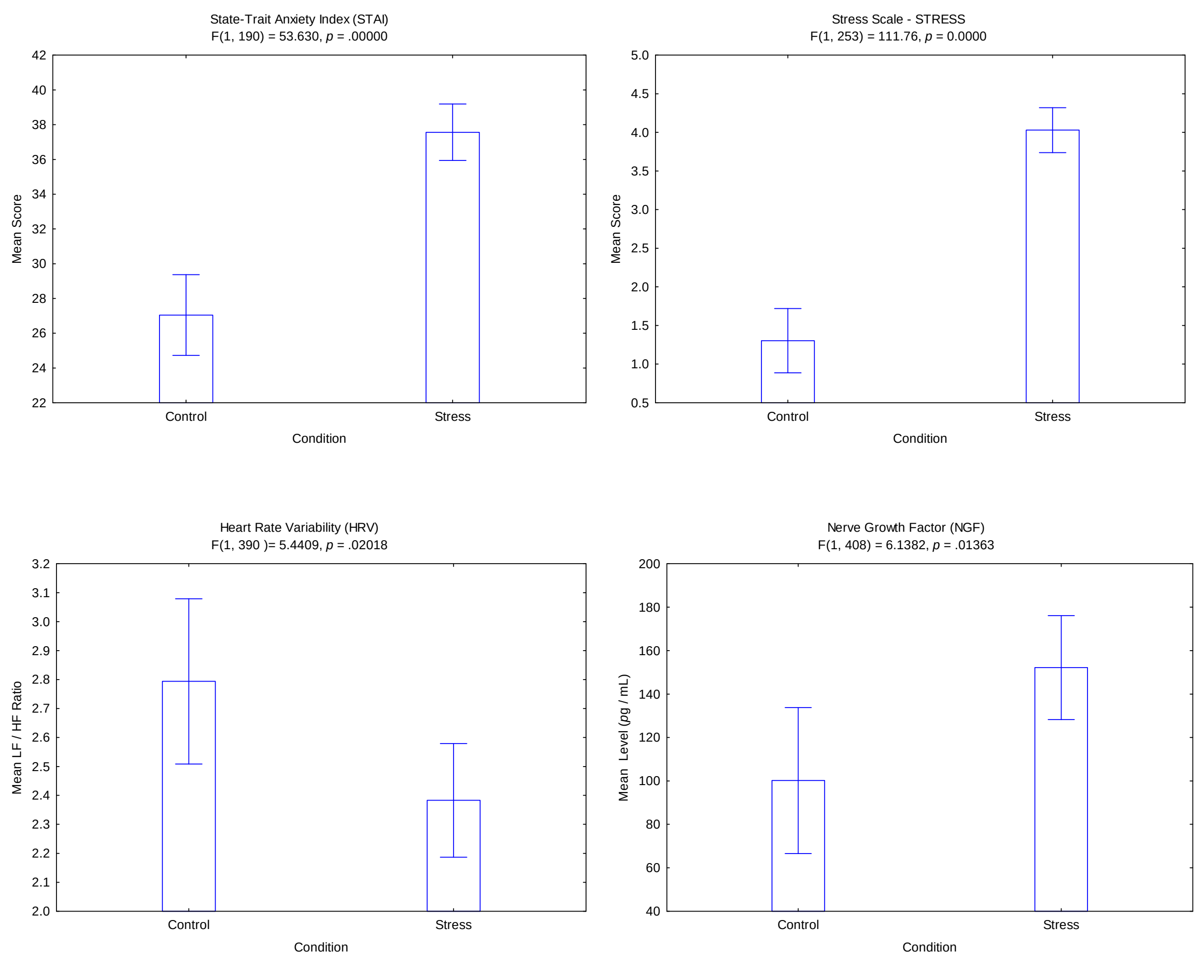

Figure 5. ANOVA Results – Group Comparisons.

OBJECTIVE B – EVALUATING THE IMPACT OF DECEPTION

The next task was to determine whether experimenter interaction (i.e., Deception) was a significant amplifier of the stress response. This was achieved by comparing the subjects who had only received the stressor task (Stress group) with those who had also received experimenter interaction during and after the stressor task (Stress + Deception).

Although a series of one-way ANOVA tests failed to reach significance for any measure, it should be noted that most of the results were in the direction of confirming the effectiveness of the Deception manipulation, and some measures were close to statistical significance, e.g., STAI; $F(1, 127) = 6.5961$, $p = .01138$. Nevertheless, these results were taken, as a whole, to indicate that Deception did not significantly add to the level of the stress response elicited by either task.

OBJECTIVE C – COMPARING STRESS TASKS

Both the stress response and stress measurement tools were next evaluated by comparing results of Control subjects and Stress (only) subjects for each computer task, and then comparing corresponding Stress groups for both tasks. MIST results are shown in Figure 6 and VBS2 results are shown in Figure 7. STAI, STRESS, and NGF (either as raw data or as a natural logarithm transform) were significant for both tests.

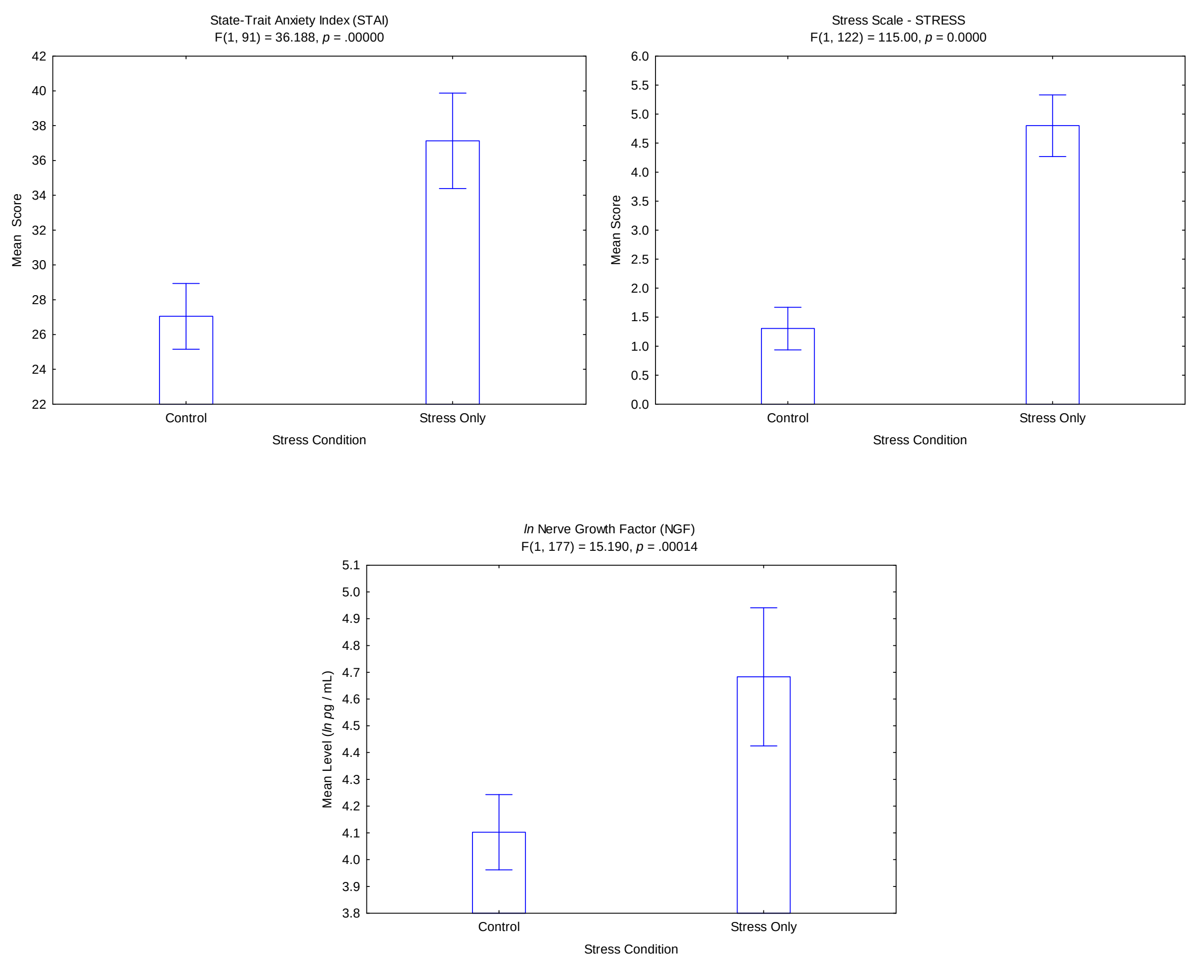

Figure 6. ANOVA Results – MIST Task.

The figures show evidence for stress response based primarily on psychological tools. While other measures were largely in the expected direction, none reached the threshold for statistical significance. Clearly, the effect size was greater with the larger sample sizes involved in comparisons of combined tasks (Figure 5).

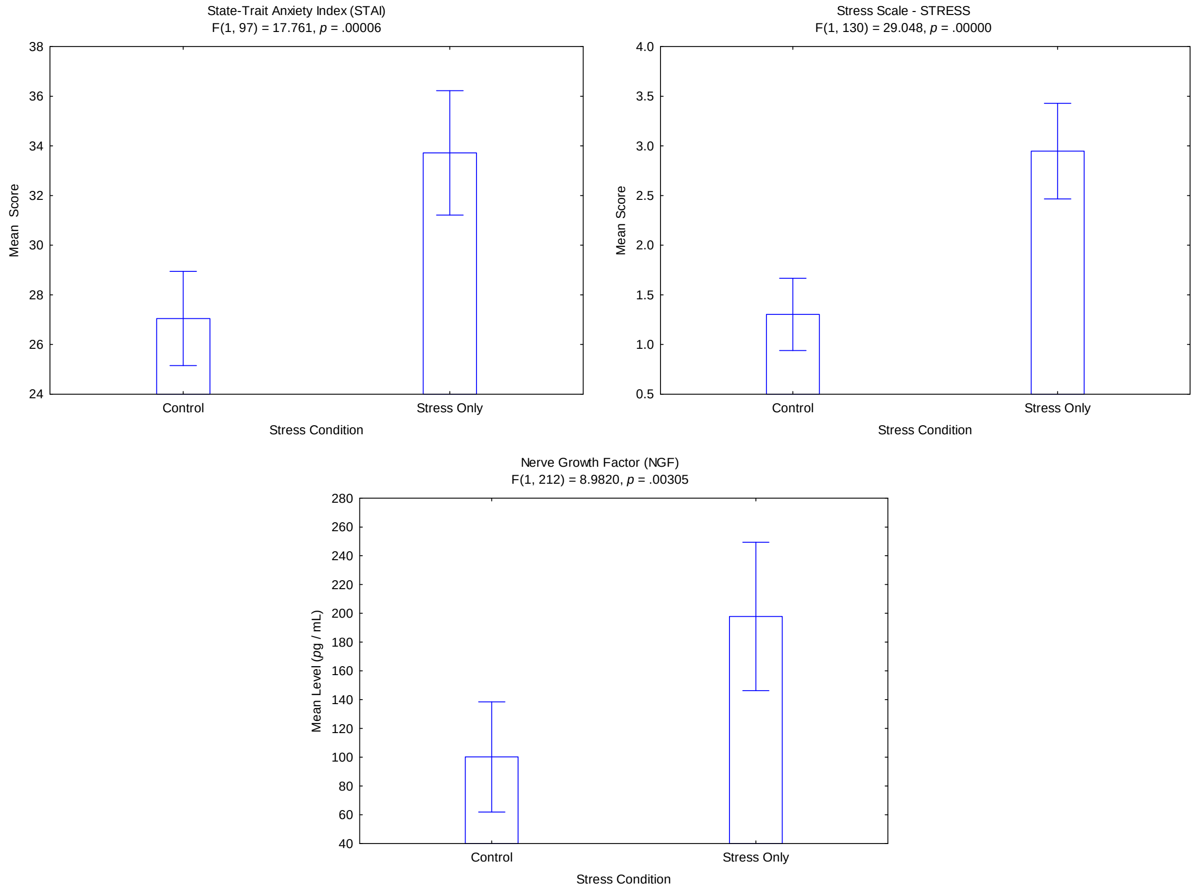

Figure 7. ANOVA Results – VBS2 Task.

A direct comparison of stress response between the MIST and VBS2 tasks is shown in Figure 8, where only the STRESS scale reached significance.

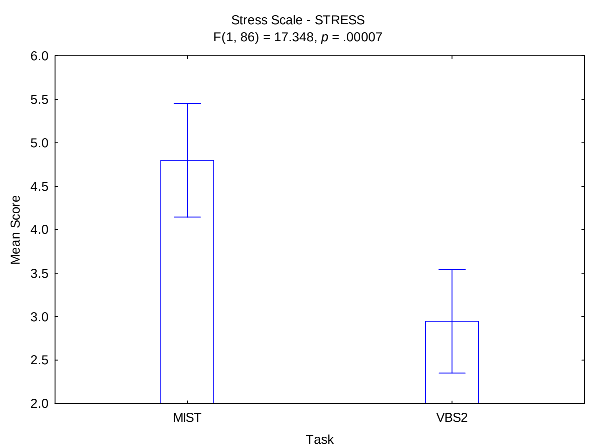

Figure 8. ANOVA Results – MIST – VBS2 Task Comparison.

Based only on the sparse results of the task comparison (Figure 8), a meaningful distinction between the effectiveness of the two stressor tasks cannot be made.

OBJECTIVE D – THE IMPACT OF STRESS ON COGNITIVE PERFORMANCE

Speed (RT) and accuracy (Percent Correct) results of the ANAM test series were analyzed separately. Neither ANOVA test was significant. Direct inspection of raw results revealed performance changes in both directions (i.e., improvement or decrement) between Control and Stress conditions, with high variability in all measures.

Considering that subject stress levels were likely to be higher immediately following completion of the stress task, the initial ANAM task of the nine-task series—Code Substitution—was examined separately, but this analysis also failed to reach significance.

These results indicate that the stress levels generated by the computer tasks either had no impact on subsequent cognitive performance or that the levels were of insufficient intensity to achieve an impact.

OBJECTIVE E – EVALUATING STRESS RESPONSE MEASURES

Both psychological instruments—STAI and STRESS—provided consistent and interpretable measures of stress response at different levels of analysis (Figures 5–8). With the exception of NGF, however, none of the physiological measures demonstrated consistent performance. HRV was significant and interpretable only for the overall data set (i.e., involving both MIST and VBS2 results), while HR and cortisol showed complex patterns and little statistical significance.

In summary, these results indicate that STAI or STRESS, and NGF, represent the most promising performance measures for future phases of this research.

DISCUSSION

Although the study was successful in generating a stress response, the performance of individual response measures was not uniformly sensitive or consistent. While the STAI and STRESS instruments yielded good results as response measures, it remains to be seen whether psychological tools are both necessary and sufficient for stress research, especially in light of the generally satisfactory results obtained in other research with a similar variety of physiological tools. Most troubling was the performance of the heart rate and cortisol measures.

HEART RATE

HR, like all other cardiac measures, was analyzed using spectral methods. The Aria system required special software to open the data records and, while a delmar Reynolds data reduction program is available for analysis, the decision to perform our own research analysis was influenced by the significant success enjoyed by the Stony Brook Biomedical Research Department using specially developed algorithms. While the spectral channels examined with these algorithms (e.g., low frequency, high frequency, sympathetic and parasympathetic power spectral densities) provided excellent results, the heart rate metric might have been better examined using time domain methods and a shorter (e.g., 1-minute) sampling interval. The Stony Brook algorithm capability is, however, not currently available. Furthermore, the data reduction method would not seem to explain the divergent HR values between the Control and the Stress groups outside of the *Training—ANAM* interval. For these reasons, resolution of this issue will most likely have to wait until a new research design is executed.

HEART RATE VARIABILITY

Where physiological data reduction processes returned missing or suspect results, the data were excluded from analysis, which meant that some statistical tests involved fewer subjects than others. This procedure might have affected HRV results, owing to small sample sizes. Therefore, a power analysis was performed of the HRV data to determine whether the number of subjects might have impacted the results, using a one-tailed *t*-test for independent means. Raw data for these HRV conditions, and the resulting statistical power, are shown in Table 4.

Based on this analysis, 70+ subjects would be required for each task (MIST or VBS2) to obtain a power of 0.8—reasonable for research purposes (e.g., Statsoft, 2010)—and to resolve HRV at the task level.

Table 4. Non-significant HRV Results.

ANOVA Results	MIST	VBS2
	F(1, 187) = 2.5526, *p* = .11180	F(1, 201) = 0.99844, *p* = .31889

MIST	**Means**		**Valid Subject *n***		**Power**
	Control	Stress Only	Control	Stress Only	0.293
	2.68717	1.823884	18	9	

VBS2	**Means**		**Valid Subject *n***		**Power**
	Control	Stress Only	Control	Stress Only	0.181
	2.68717	2.48664	18	11	

CORTISOL

Earlier stress research provided critical guidance to the study reported here. Common to these studies is a salivary cortisol profile that elevates following a stressor, and then returns approximately to a pre-stress baseline, while cortisol profiles for matched control groups remain relatively flat. The effect is not, however, universal. Figure 9 depicts salivary cortisol patterns for the current study together with those of Dedovic, et al. (2005) and Wang, et al., (2005). Values have been standardized as ratios of the first (baseline) sample for each experiment (as Dedovic reports cortisol in unitis of nmol / L). Furthermore, test points are approximate and based on experimenter judgment, as the three protocols were not identical.

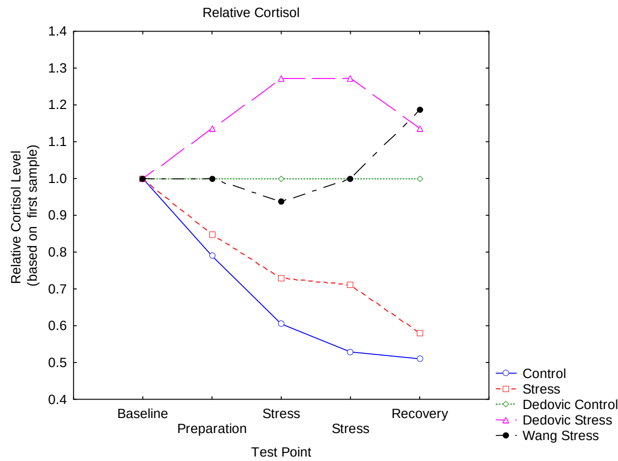

Figure 9. Cortisol Level Comparison (Dedovic, et al., 2005 and Wang, et al., 2005).

The Dedovic pattern is most common in the literature reviewed for this study. The Wang result, i.e., a significant (up to 24-minute) delay in stress response, serves to illustrate the range of profiles possible with salivary cortisol. By way of comparison, Figures 3 and 4 show that the cortisol profiles are not entirely unreasonable during the *Training* to *ANAM 1* interval—a period of approximately 25 minutes.

The current cortisol measure may, therefore, represent an interpretable distinction between the Control and Stress groups for the period following *Training*. More difficult to explain, however, is the steady negative slope of both the Stress and Control profiles throughout the experiment session which remains unresolved.

COGNITIVE PERFORMANCE

Ensuring that a stress response could be elicited using a computer task, and evaluating alternative response measures, were two critical objectives of the study. The primary objective, however, was to determine the impact of that stress on cognitive performance, as an antecedent step toward comparing both stress and fatigue effects within the same experiment paradigm; Failure to identify such effects at this level makes any attempt at more refined research pointless. Three possible explanations arise to account for this lack of results.

The research reported here relied on a between-subjects design, i.e., comparing cognitive performance across groups that had differed in their exposure to a stress task. This decision was made to avoid the “learning effects” problem of within-subjects designs. The within-subjects approach allows each subject to serve as their own control by exposing every subject to every element of the experiment. Applying this approach to the current study would require each subject to complete the ANAM battery twice (both before and after the stress task), to evaluate performance differences. The process, however, would simultaneously provide subjects with an opportunity to gain skill with these tasks, which could bias results. An additional penalty of within-subjects designs is cost; this approach, while more sensitive, would require subjects to return repeatedly to the test facility over a period of days, adding significant cost in staffing, time commitment and subject payment. In addition, there was great concern regarding subject attrition (i.e., subjects not returning to complete all phases of the study) that influenced the final design choice.

A second reason for the current design was to provide an opportunity to isolate the stress effects of completing the ANAM battery itself. This was accomplished via the Control group and, as seen in the STAI and STRESS profiles of Figure 5, a measurable effect did exist. Clearly, however, design decisions had an impact on the sensitivity to performance changes. Such research decisions always have consequences that must be identified and traded off against each other. Based on these results, a within-subjects design might have proven to be the better choice.

A third perspective on cognitive performance is that the use of other measurement tools might have provided improved results. As discussed earlier, previous research has shown that stress and fatigue demonstrate a variety of both congruent and non-congruent performance effects. The design executed here was influenced by the need to measure a wide range of cognitive effects as a foundation for further research that would involve both stress and fatigue, and the ANAM battery appeared to be the best tool for covering this diverse territory. Selecting for comprehensive measurement, however, necessarily led to a lengthy testing process involving two sessions of ANAM tests. As seen in Figure 5, stress levels deteriorated during this period, indicating that the stress response was different toward the end of the ANAM series than it was at the beginning.

STRESS TASKS

The intensity of the stressor tasks was established with a view toward Institutional Review Board (IRB) concerns; this initial foray into human stress research generated considerable discussion among IRB members regarding the extent of the stress manipulation, and both MIST and VBS2 represented a consensus of all parties regarding research objectives and subject protection. The deception condition, included to ensure sufficient stimulus intensity to evoke a stress response, was also designed for exploration only, with no real intention to use such measures in future phases of the work unless absolutely necessary.

It is reasonable to believe that military stress research is most useful for tasks that are most relevant to military performance. In this perspective, the VBS2 stressor task would appear to have many advantages in stress research involving a military audience. Unfortunately, the distinction between the two stress tasks is unresolved. Only the STRESS measure yielded a significant distinction between MIST and VBS2 which does not, by itself, represent a compelling reason to conclude that these tasks are different. All other results were not significant, and even an inspection of the raw data showed equivocal patterns.

These results compelled a further review of the literature, to better illuminate the mechanisms behind human stress response. In fact, any task is stressful to the extent that the subject perceives a mismatch between the demands of a task and their ability to cope with those demands (e.g., Lazarus and Folkman, 1984); the greater the mismatch, the greater the stress. Individual stress response is, therefore, an outcome of personal judgment or appraisal of task demands (e.g., Matthews, 2003) and any approach to characterizing stress must account for the individual psychological factors that enter into that response.

We therefore approach the selection of a stressor task with considerable care, and propose that the next phase of this research focus on potential differences in task appraisal, using an extended experiment regarding the VBS2 task. Specifically, we propose a study of this combat task by comparing the responses of two populations that differ only in their exposure to the combat environment. The outcome of that work will determine which task—MIST or VBS2—will be used in future work.

CONCLUSIONS

The experiment succeeded in generating a measurable stress response in the laboratory, using two different task manipulations, achieving a critical objective of the study. No impact was found, however, on cognitive performance as a result of stress, possibly due to the intensity of the stressor, the choice of the cognitive test battery, the duration of the test battery, or the between-subjects experiment design selected for the study. The choice of stressor task for further phases of this work is deferred, pending an interim study to examine aspects of human stress appraisal; this step is necessary to ensure a complete understanding of task characteristics that may bear on further stress testing.

The effectiveness and interpretation of some stress response measures was not completely resolved. While several tools—psychological instruments and NGF—proved useful, other measures did not perform consistently or did not perform as expected. Methodological issues appear to be the primary cause of these anomalies, which can be corrected in future work.

REFERENCES

Aloe, L., Alleva, E., and Fiore, M. (2002). Stress and nerve growth factor: Findings in animal models and humans. *Pharmacology Biochemistry and Behavior*, 73(1) 159-166.

Aloe, L., Alleva, E., Bohm, A., and Levi-Montalcini, R. (1986). Aggressive behavior induces release of nerve growth factor from mouse salivary gland into the bloodstream. *Proceedings of the National Academy of Sciences, 83(16).*

Bohemia Interactive (2010) http://www.bistudio.com/bohemia-interactive-simulations/virtual-battlespace-2_en.html

Borg, G.A.V. (1982). Psychophysical bases of perceived exertion. Medicine and Science in Sports and Exercise, 14(5).

Box, G.E., Hunter, W.G., and Hunter, J.S. (1978). Statistics for Experimenters. Hoboken, NJ: John Wiley and Sons.

Brisswalter, J., Collardeau, M., and Rene, Arcelin. (2002). Effects of acute physical exercise characteristics on cognitive performance. *Sports Medicine*, 32 (9).

Cardena, E., Koopman, C., Classen, C., Waelde, L.C., and Spiegel, D. (2000). Psychometric properties of the Stanford Acute Stress Reaction Questionnaire (SASRQ): A valid and reliable measure of acute stress. *Journal of Traumatic Stress*, 13(4).

Chiffer McKay, K.A., Buen, J.E, Bohan, K.J., and Maye, J.P. (2010). Determining the relationship of acute stress, anxiety, and salivary a-amylase level with performance of student nurse anesthesiologists during human-based anesthesia simulator training. *AANA Journal, 78(4).*

C-SHOP: Center for the Study of Human Operator Performance (2010). http://c-shop.ou.edu/anam4.htm

Davidson, M.J. and Cooper, C.L. (1981). A model of occupational stress. *Journal of Occupational Medicine*, *23(8*) 564–574.

Dedovic, K., Renwick, R., Mahani, N.K., Engert, V., Lupien, S.J., and Pruessner, J.C. (2005). The Montreal Imaging Stress Task: Using functional imaging to investigate the effects of perceiving and processing psychosocial stress. Journal of Psychiatry and Neuroscience, 30(5).

Filaire, E., Portier,H., Massart, A., Ramat, L., and Teixeira, A. (2009). Effect of lecturing to 200 students on heart rate variability and alpha-amylase. *European Journal of Applied Physiology*, 108(5).

Friedl, K.E., Mallis, M.M., Ahlers, S.T., Popkin, S.M., and Larkin, W. (2004). Research requirements for operational decision-making using models of fatigue and performance. *Aviation, Space, and Environmental Medicine*, *75(3*).

Gorman, J.C., Cooke, N.J., and Winner, J.L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics*, 49 (12-13), 1312-1325.

Grossman, P. (1983). Respiration, stress, and cardiovascular function. *Psychophysiology*, 20(3).

Gunzelmann, G., Kershner, J., Van Dongen, H.P.A., and Dinges, D.F. (2007). Understanding decrements in knowledge access resulting from increased fatigue. In D.S. McNamara and J.G. Trafton (Eds.), *Proceedings of the Twenty-Ninght Annual Meeting of the Cognitive Science Society*. Austin, TX.

Habeck, C., Rakitin, B.C., Moeller, J.,Scarmeas, N., Zarahn, E., Brown,T., and Stern Y. (2004). An event-related fMRI study of the neurobehavioral impact of sleep deprivation on performance of a delayed-match-to-sample task. *Cognitive Brain Research*, 18, 306-321.

Harrison, Y. and Horne, J.A. (2000). The imact of sleep deprivation on decision making: A review. *Journal of Experimental Psychology: Applied. 6(3*), 236-249.

Horvath, F. (1978). An experimental comparison of the psychological stress evaluator and the galvanic skin response in detection and deception. *Journal of Applied Psychology*, 63(3).

Huskisson E.C. (1982). Measurement of pain. *Journal of Rheumatology*, 9(5)

Joels, M., Pu, Z., Wiegert, Oitz, M.S., and Krugers, H.J. (2006). Learning under stress: how does it work? In Trends in Cognitive Sciences, 10:152-158.

Keinan, G. (1987). Decision making under stress: Scanning of alternatives under controllable and uncontrollable threats. *Journal of Personality and Social Psychology*, 52(3).

Kirschbaum, C, Klauer,T., Filipp, S.H., and Hellhammer, D.H. (1995). Sex-specific effects of social support on cortisol and subjective responses to acute psychological stress. *Psychosomatic Medicine*, 57(1)

Kirschbaum, C. and Hellhammer, D.H. (1999). Noise and stress – salivary cortisol as a non-invasive measure of allostatic load. Noise Health (1) 57–65

Kirschbaum, C. Wolf, O.T., May, M., Wippich, W. and Hellhammer, D.H. (1996). Stress- and treatment-induced elevations of cortisol levels associated with impaired declarative memory in healthy adults. *Life Sciences*, 58(17).

Kirschbaum, C., Pirke, K-M., and Hellhammer, D.H. (1993). The 'Trier Social Stress Test' – A tool for investigating psychobiological stress responses in the laboratory. *Neuropsychobiology*, 28: 76–81.

Krampen, G. (1991). Fragebozen zu Kompetenz- und Kontrolluberzeugungen (FKK) Gottingen: Hogrefe.

Krueger, G.P. (1989). Sustained work, fatigue, sleep loss and performance: A review of the issues. *Work and Stress*, 3(2).

Lazarus, R.S., and Folkman, S. (1984). Stress, Appraisal and Coping. New York: Springer Publishing Company.

Leach, J. (2004). Why people "freeze" in an emergency: Temporal and cognitive constraints on survival responses. Aviation, Space, and Environmental Medicine, 75 (539-542).

Lieberman, H.R., Bathalon, G.P., Falco, C.M., Morgan III, C.A. Niro, P.J., and Tharion, W.J. (2005).

McGonagle, K.A. and Kessler, R.C. (1990). Chronic stress, acute stress, and depressive symptoms. *American Journal of Community Psychology*, 18(5).

McNair, D.M, Lorr, M., and Droppelman, L.F. (1971). Manual for the profile of mood states. San Diego, CA: Educational and Industrial Testing Service.

Medcompare (2010). http://www.medcompare.com/details/3760/AriaAndreg-Digital-Holter-Recorder.html

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., and Hodges, J.R. (2006). The Addenbrooke's cognitive examination revised (ACE-R): A cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry*, 21(11)

Noto, Y., Sato, T., Kudo, M., Kurata, K., and Hirota, K. (2005). The relationship between salivary biomarkers and state-trait anxiety inventory score under mental arithmetic stress: A pilot study. *Anesthesiology and Analgesia*, *101* (1873–6).

Orosanau, J.M., and Backer, P. (1996). Stress and military performance. In J.E. Driskell & E. Salas (Eds.), Stress and Human Performance. Mahwah (NJ): Lawrence Erlbaum.

Poteliakhoff, A. (1981). Adrenocortical activity and some clinical findings in acute and chronic fatigue. *Journal of Psychosomatic Research*, 25(2).

Raggatt, P.T.F and Morrissey, S.A. (1997). A field study of stress and fatigue in long-distance bus drivers. *Behavioral Medicine, 23 (3*). 122-129.

Reeves, D.L. and Winter, K.P. (1992). ANAM documentation volume 1: Test administrator's guide. Office of Military Performance Technology, Scienceboard, Washington, DC

Rosekind, M.R., Philippa, H, Miller, D.L., Gregory, K.B., Smith, R.M., Weldon, K.J., Co, E.L., McNally, K.L., and Lebacqz, J.V.(1994). Fatigue in operational settings: Examples from the aviation environment. *Human Factors*, 36(2).

Rosenberg, M. (1979). Conceiving the Self. New York: Basic Books.

Spielberger, C.D. and Sydeman, S.J. (1994). State-Trait Anxiety Inventory and State-Trait Anger Expression Inventory. . In The Use of Psychological Testing for Treatment Planning and Outcome Assessment (Ch 13) 292-321. Hillsdale, NJ: Erlbaum Associates.

Statsoft (2010). http://www.statsoft.com/ .

Steptoe, A., Hamer, M., and Chida, Y. (2007). The effects of acute psychological stress on circulating inflammatory factors in humans: A review and meta-analysis. *Brain, Behavior, and Immunity, 21(7).*

Sterling, B.S. and Perala, C.H (2007). Workload, stress, and situation awareness of soldiers who are controlling unmanned vehicles in future urban operations. Army Research Laboratory (ARL-TR-4071), April.

Thorne, D.R., Johnson, D.E., Redmond, D.P., Sing, H.C., and Belenky, G. (2005). *Behavior Research Methods*, 37.

U.S. Army. (1998) Combat Stress Control in a Theater of Operations. C1, FM 8-51. Headquarter Department of the Army Washington, DC, 30 January 1998.

Van Dongen, H.P., Baynard, M.D., Maislin, G., and Dinges, D.F. (2004). Systematic interindividual differences in neurobehavioral impairment from sleep loss: Evidence of trait-like differential vulnerability. *Sleep*, 27(3).

Van Dongen, H.P.A., Maislin, G., Mullington, J.M., and Dinges, D.F. (2003). The cumulative cost of additional wakefulness: Does-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep, 26(2*).

Vedhara, K., Hyde, J., Gilchrist, I.D., Tytherleigh, M., and Plummer, S. (2000). Acute stress, memory, attention, and cortisol. *Psychoneuroendocrinology*, 25(6).

Vrijkotte, T.G.M., van Doornen, L.J.P., and de Geus, E.J.C. (2000). Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability. Hypertension, 35(880).

Wang, Z., Rao, H., Wetmore, G.S., Furlan, P.M., Korczykowski, M., Dinges, D.F., and Detre, J.A. (2005). Perfusion functional MRI reveals cerebral blood flow pattern under psychological stress. Proceedings of the national Academy of Sciences, 102(49), 17804-17809.

Yehuda, R., Keefe, R.S., Harvey, P.D., Levengood, R.A., Gerber, D.K., Geni, J.,and Siever, L.J. (1995). Learning and memory in combat veterans with posttraumatic stress disorder. *American journal of Psychiatry*, (152).

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-01-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
December 2010	Final	

4. TITLE AND SUBTITLE

The Effects of Acute Stress on Cognitive Performance
A Pilot Study

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHORS

Steven A. Murray, Matthew Yanagi, Wayne Ensign
SSC Pacific

Camellia Clark, Burcu Darst
University of California, San Diego

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

SSC Pacific
San Diego, CA 92152–5001

8. PERFORMING ORGANIZATION REPORT NUMBER

TR 1995

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)

ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

A laboratory study was conducted to establish a research framework for investigating the effects of stress and fatigue on cognitive performance. The initial objectives were to (a) confirm the effectiveness of candidate stressor tasks, (b) to evaluate alternative stress response measures, and (c) to benchmark a series of cognitive performance tests. Both stressor tasks proved effective in eliciting stress under laboratory conditions, as indicated by multiple stress measures. All measures, however, were not effective, and no cognitive performance effects were found. Results are explained in terms of experiment design factors (i.e., the between-subjects approach used for the study) and the intensity and duration of stress levels achievable under laboratory conditions. Methodological revisions and an interim experiment are discussed in the context of larger research objectives that address both stress and fatigue together.

15. SUBJECT TERMS

Mission Area: Research and Applied Sciences
Cognitive performance
Stress response measures

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Steven A. Murray
					19B. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	36	(619) 553–6350

INITIAL DISTRIBUTION

85300	Archive/Stock	(1)
71730	S. Murray	(1)

Defense Technical Information Center
Fort Belvoir, VA 22060–6218 (1)

Approved for public release; distribution is unlimited.

.

SSC Pacific
San Diego, CA 92152-5001